**Description**

`pkequiv` performs bioequivalence testing for two treatments. By default, `pkequiv` calculates a standard confidence interval (CI) symmetric about the difference between the two treatment means. `pkequiv` also calculates CIs symmetric about 0 and intervals based on Fieller’s theorem. Also, `pkequiv` can perform interval hypothesis tests for bioequivalence.

`pkequiv` is one of the `pk` commands. Please read `[R] pk` before reading this entry.

**Quick start**

Classic CI for difference in pharmacokinetic outcome $y_1$ between treatments $v_1$ given over period $v_2$ in sequence $v_3$ with subjects identified by $idvar$

```
pkequiv y1 v1 v2 v3 idvar
```

As above, but calculate an exact CI by Fieller’s theorem

```
pkequiv y1 v1 v2 v3 idvar, fieller
```

Schuirmann’s two one-sided tests for bioequivalence

```
pkequiv y1 v1 v2 v3 idvar, tost
```

Specify the two treatments, 2 and 3, that are to be tested for equivalence

```
pkequiv y1 v1 v2 v3 idvar, compare(2 3)
```

**Menu**

Statistics > Epidemiology and related > Other > Bioequivalence tests
Syntax

pkequiv outcome treatment period sequence id [if] [in] [ , options ]

<table>
<thead>
<tr>
<th>options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>compare(string)</td>
<td>compare the two specified values of the treatment variable</td>
</tr>
<tr>
<td>limit(#)</td>
<td>equivalence limit (between 0.01 and 0.99); default is 0.2</td>
</tr>
<tr>
<td>level(#)</td>
<td>set confidence level; default is level(90)</td>
</tr>
<tr>
<td>fieller</td>
<td>calculate CI by Fieller’s theorem</td>
</tr>
<tr>
<td>symmetric</td>
<td>calculate symmetric CI</td>
</tr>
<tr>
<td>anderson</td>
<td>Anderson and Hauck hypothesis test for bioequivalence</td>
</tr>
<tr>
<td>tost</td>
<td>two one-sided hypothesis tests for bioequivalence</td>
</tr>
<tr>
<td>noboot</td>
<td>do not estimate probability that CI lies within confidence limits</td>
</tr>
</tbody>
</table>

Options

compare(string) specifies the two treatments to be tested for equivalence. Sometimes there may be more than two treatments, but the equivalence can be determined only between any two treatments.

limit(#) specifies the equivalence limit. The default is 0.2. The equivalence limit can be changed only symmetrically; for example, it is not possible to have a 0.15 lower limit and a 0.2 upper limit in the same test.

level(#) specifies the confidence level, as a percentage, for CIs. The default is level(90). This setting is not controlled by the set level command.

fieller specifies that a CI based on Fieller’s theorem be calculated.

symmetric specifies that a symmetric CI be calculated.

anderson specifies that the Anderson and Hauck (1983) hypothesis test for bioequivalence be computed. This option is ignored when calculating a CI based on Fieller’s theorem or when calculating a CI that is symmetric about 0.

tost specifies that the two one-sided hypothesis tests for bioequivalence be computed. This option is ignored when calculating a CI based on Fieller’s theorem or when calculating a CI that is symmetric about 0.

noboot prevents the estimation of the probability that the CI lies within the equivalence limits. If this option is not specified, this probability is estimated by resampling the data.

Remarks and examples

pkequiv is designed to conduct tests for bioequivalence based on data from a crossover experiment. pkequiv requires that the user specify the outcome, treatment, period, sequence, and id variables. The data must be in the same format as that produced by pkshape; see [R] pkshape.
We use `pkdata.dta` from example 2 of [R] `pk`. We use `pkcollapse` and `pkshape` on the data as discussed in that example. After sorting the data with `sort`, our data appear as follows:

```stata
use https://www.stata-press.com/data/r16/pkdata, clear
pkcollapse time conc1 conc2, id(id) keep(seq) stat(auc)
........................
pkshape id seq auc*, order(RT TR)
.sort period id
.list, sep(4)
```

<table>
<thead>
<tr>
<th>id</th>
<th>sequence</th>
<th>outcome</th>
<th>treat</th>
<th>carry</th>
<th>period</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1</td>
<td>RT</td>
<td>R</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>2</td>
<td>RT</td>
<td>R</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3.</td>
<td>3</td>
<td>RT</td>
<td>R</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4.</td>
<td>4</td>
<td>RT</td>
<td>R</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5.</td>
<td>5</td>
<td>RT</td>
<td>R</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>6.</td>
<td>6</td>
<td>RT</td>
<td>R</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>7.</td>
<td>7</td>
<td>RT</td>
<td>R</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>8.</td>
<td>8</td>
<td>RT</td>
<td>R</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>9.</td>
<td>9</td>
<td>TR</td>
<td>T</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>10.</td>
<td>10</td>
<td>TR</td>
<td>T</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>11.</td>
<td>11</td>
<td>TR</td>
<td>T</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>12.</td>
<td>12</td>
<td>TR</td>
<td>T</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>13.</td>
<td>13</td>
<td>TR</td>
<td>T</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>14.</td>
<td>14</td>
<td>TR</td>
<td>T</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>15.</td>
<td>15</td>
<td>TR</td>
<td>T</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>16.</td>
<td>16</td>
<td>TR</td>
<td>T</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>17.</td>
<td>1</td>
<td>RT</td>
<td>T</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>18.</td>
<td>2</td>
<td>RT</td>
<td>T</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>19.</td>
<td>3</td>
<td>RT</td>
<td>T</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>20.</td>
<td>4</td>
<td>RT</td>
<td>T</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>21.</td>
<td>5</td>
<td>RT</td>
<td>T</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>22.</td>
<td>6</td>
<td>RT</td>
<td>T</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>23.</td>
<td>7</td>
<td>RT</td>
<td>T</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>24.</td>
<td>8</td>
<td>RT</td>
<td>T</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>25.</td>
<td>9</td>
<td>TR</td>
<td>R</td>
<td>T</td>
<td>2</td>
</tr>
<tr>
<td>26.</td>
<td>10</td>
<td>TR</td>
<td>R</td>
<td>T</td>
<td>2</td>
</tr>
<tr>
<td>27.</td>
<td>11</td>
<td>TR</td>
<td>R</td>
<td>T</td>
<td>2</td>
</tr>
<tr>
<td>28.</td>
<td>12</td>
<td>TR</td>
<td>R</td>
<td>T</td>
<td>2</td>
</tr>
<tr>
<td>29.</td>
<td>13</td>
<td>TR</td>
<td>R</td>
<td>T</td>
<td>2</td>
</tr>
<tr>
<td>30.</td>
<td>14</td>
<td>TR</td>
<td>R</td>
<td>T</td>
<td>2</td>
</tr>
<tr>
<td>31.</td>
<td>15</td>
<td>TR</td>
<td>R</td>
<td>T</td>
<td>2</td>
</tr>
<tr>
<td>32.</td>
<td>16</td>
<td>TR</td>
<td>R</td>
<td>T</td>
<td>2</td>
</tr>
</tbody>
</table>
We use `pkequiv` to conduct a bioequivalence test between reference treatment “R” and test treatment “T”.

```
set seed 123
pkequiv outcome treat period seq id
```

**Classic confidence interval for bioequivalence**

<table>
<thead>
<tr>
<th></th>
<th>[equivalence limits]</th>
<th>[test limits]</th>
</tr>
</thead>
<tbody>
<tr>
<td>difference:</td>
<td>-30.296</td>
<td>30.296</td>
</tr>
<tr>
<td>ratio:</td>
<td>80%</td>
<td>120%</td>
</tr>
</tbody>
</table>

Probability test limits are within equivalence limits = 0.6590

Note: Reference treatment = 1

The default output for `pkequiv` shows a CI for the difference between the means (test limits), the ratio of the means, and the user-specified equivalence limits. The classic CI can be constructed around the difference between the average measure of effect for the two drugs or around the ratio of the average measure of effect for the two drugs. `pkequiv` reports both the difference measure and the ratio measure. Following Chow and Liu (2009), we can apply the ±20 rule to the difference and ratio measures to determine equivalence limits. For these data, the CI for the difference must be entirely contained within the range $[-30.296, 30.296]$ and for the ratio between 80% and 120%. Here the test limits are within the equivalence limits. Although the test limits are inside the equivalence limits, there is only a 66% assurance that the observed CI will be within the equivalence limits in the long run. This is an interesting case because, although this sample shows bioequivalence, the evaluation of the long-run performance indicates possible problems. These fictitious data were generated with high intrasubject variability, which causes poor long-run performance.

We now use the data published in Chow and Liu (2009, 71), which we describe in example 1 of [R] `pkshape`. As shown in example 1, we also use `pkshape` to reshape the data. We use the same `pkequiv` command used above to conduct a bioequivalence test on the data.

```
use https://www.stata-press.com/data/r16/chowliu, clear
pkshape id seq period1 period2, order(RT TR)
set seed 123
pkequiv outcome treat period seq id
```

**Classic confidence interval for bioequivalence**

<table>
<thead>
<tr>
<th></th>
<th>[equivalence limits]</th>
<th>[test limits]</th>
</tr>
</thead>
<tbody>
<tr>
<td>difference:</td>
<td>-16.512</td>
<td>16.512</td>
</tr>
<tr>
<td>ratio:</td>
<td>80%</td>
<td>120%</td>
</tr>
</tbody>
</table>

Probability test limits are within equivalence limits = 0.9940

Note: Reference treatment = 1

For these data, the test limits are well within the equivalence limits, and the probability that the test limits are within the equivalence limits is 99.4%. 

---

1
Example 2

pkequiv displays interval hypothesis tests of bioequivalence if you specify the tost or the anderson option, or both. For example,

\begin{verbatim}
   . set seed 123
   . pkequiv outcome treat period seq id, tost anderson

   Classic confidence interval for bioequivalence

<table>
<thead>
<tr>
<th>[equivalence limits]</th>
<th>[ test limits ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>difference:</td>
<td></td>
</tr>
<tr>
<td>-16.512</td>
<td>16.512</td>
</tr>
<tr>
<td>80%</td>
<td>120%</td>
</tr>
<tr>
<td>-8.698</td>
<td>4.123</td>
</tr>
<tr>
<td>89.464%</td>
<td>104.994%</td>
</tr>
</tbody>
</table>

   Probability test limits are within equivalence limits = 0.9940

   Schuirmann’s two one-sided tests

   upper test statistic = -5.036 p-value = 0.000
   lower test statistic = 3.810  p-value = 0.000

   Anderson and Hauck’s test

   noncentrality parameter = 4.423
   test statistic = -0.613  empirical p-value = 0.0005

Note: Reference treatment = 1
\end{verbatim}

Both of Schuirmann’s one-sided tests are highly significant, suggesting that the two drugs are bioequivalent. A similar conclusion is drawn from the Anderson and Hauck test of bioequivalence.

Stored results

pkequiv stores the following in r():

Scalars

r(stddev)         pooled-sample standard deviation of period differences from both sequences
r(uci)            upper limit of a classic CI
r(lci)            lower limit of a classic CI
r(delta)          delta value used in calculating a symmetric CI
r(u3)             upper limit of Fieller’s CI
r(l3)             lower limit of Fieller’s CI

Methods and formulas

The lower limit for the difference in the two treatments for the classic shortest CI is

\[ L_1 = (\bar{Y}_T - \bar{Y}_R) - t(\alpha, n_1 + n_2 - 2) \hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \]

The upper limit is

\[ U_1 = (\bar{Y}_T - \bar{Y}_R) + t(\alpha, n_1 + n_2 - 2) \hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \]
The limits for the ratio measure are

\[ L_2 = \left( \frac{L_1}{Y_R} + 1 \right) 100\% \]

and

\[ U_2 = \left( \frac{U_1}{Y_R} + 1 \right) 100\% \]

where \( \bar{Y}_T \) is the marginal mean of the test formulation of the drug, \( \bar{Y}_R \) is the marginal mean of the reference formulation of the drug, and \( t_{(\alpha,n_1+n_2-2)} \) is the \( t \) distribution with \( n_1 + n_2 - 2 \) degrees of freedom. \( \hat{\sigma}_d \) is the pooled sample variance of the period differences from both sequences, defined as

\[
\hat{\sigma}_d = \frac{1}{n_1 + n_2 - 2} \sum_{k=1}^{2} \sum_{i=1}^{n_k} (d_{ik} - \bar{d}_k)^2
\]

The finite sample performance of the classical CI is assessed via bootstrap simulation of the CI. One thousand bootstrap samples are drawn using the patient IDs as clusters. For each sample, the classical CI is constructed and compared with the equivalence limits.

The upper and lower limits for the symmetric CI are \( \bar{Y}_R + \Delta \) and \( \bar{Y}_R - \Delta \), where

\[
\Delta = k_1 \hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2} - (\bar{Y}_R - \bar{Y}_T)}
\]

and (simultaneously)

\[
\Delta = -k_2 \hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2} + (\bar{Y}_R - \bar{Y}_T)}
\]

and \( k_1 \) and \( k_2 \) are computed iteratively to satisfy the above equalities and the condition

\[
\int_{k_1}^{k_2} f(t) dt = 1 - 2\alpha
\]

where \( f(t) \) is the probability density function of the \( t \) distribution with \( n_1 + n_2 - 2 \) degrees of freedom.

See Chow and Liu (2009, 88–92) for details about calculating the CI based on Fieller’s theorem.

The two test statistics for the two one-sided tests of equivalence are

\[
T_L = \frac{(\bar{Y}_T - \bar{Y}_R) - \theta_L}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}
\]

and

\[
T_U = \frac{(\bar{Y}_T - \bar{Y}_R) - \theta_U}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}
\]

where \(-\theta_L = \theta_U\), both of these being the regulated confidence limits.
The logic of the Anderson and Hauck test is tricky; see Chow and Liu (2009) for a complete explanation. However, the test statistic is

\[ T_{AH} = \frac{(\bar{Y}_T - \bar{Y}_R) - (\theta_L + \theta_U)}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \]

and the noncentrality parameter is estimated by

\[ \hat{\delta} = \frac{\theta_U - \theta_L}{2\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \]

The empirical p-value is calculated as

\[ p = F_t \left( |T_{AH} - \hat{\delta}| \right) - F_t \left( -|T_{AH} - \hat{\delta}| \right) \]

where \( F_t \) is the cumulative distribution function of the t distribution with \( n_1 + n_2 - 2 \) degrees of freedom.

References


Also see

[R] pk — Pharmacokinetic (biopharmaceutical) data